

京 都 光 華 女 子 大 学
研 究 紀 要 第 45 号 抜 刷
平 成 19 年 12 月 5 日 発 行

社会科学データ分析のための 線形混合モデル入門

土 居 淳 子

社会科学データ分析のための 線形混合モデル入門

土 居 淳 子

1 はじめに

社会科学データの多くは、何らかの階層構造あるいはクラスター構造を持つことが多い。たとえば、同じ親をもつ子ども同士は、大きな母集団から無作為に選ばれた他人同士よりも、身体的・精神的に似通っている可能性が高いという意味で、また生活環境を共有しているという意味で1つのクラスターを構成している。また、学校に所属する生徒は、生徒が所属するクラスから影響を受け、さらにそのクラスは学校からも影響を受ける。加えて、生徒や学校は、その居住地や所在地など地域の影響も受ける。つまり、生徒<クラス<学校<地域という階層構造（入れ子構造）がある。一般に、個人は集団や社会的ネットワークの入れ子構造のなかに存在し、さまざまな社会的文脈と相互作用している。したがって、人を最小の分析単位とする社会科学データは、複雑に入り組んだ社会的文脈のなかに存在しているのである。

たとえば、市田ら（2005）が述べているように、地域の特性と地域住民の健康状態との関連を考えると、「地域住民の健康指数の平均値」と「失業率」等の地域レベルの変数間の関連を調べる分析では、地域による健康への影響を正しく分析することは出来ない。地域の失業率の高さが治安の悪化等を通して住民の健康に与える影響（地域レベルの要因）と、個人が失業して低所得であるということの影響（個人レベルの要因）の両者が混じってしまうからである。地域レベルの要因を特定するためには、個人要因と地域要因を分離し、地域要因のみを評価する必要がある。

もう一つ別の例を考えてみよう。京都市内の中学に在学する生徒に対して、数学の成績と宿題の量の関係を調べたいとする。京都市内のいくつかの中学をまず無作為に抽出し、さらにそれらの中学校の生徒から調査対象者を無作為に抽出する。一般的な分析手法として、従属(目的)変数 y を生徒の数学の成績、独立(説明)変数 x を宿題の量とする回帰モデルを考える。

$$y_i = \beta_0 + \beta_1 x_i + e_i, \quad i = 1, 2, \dots, n \quad (1)$$

$$e_i \sim N(0, \sigma^2) \quad (2)$$

ここで、添え字 i はそれぞれの生徒を表し、 n はサンプルの総数、 β_0, β_1 はデータから推定される定数、残差 e_i は平均0、分散 σ^2 をもつ正規分布からの独立なサンプルであると仮定している。

このモデルでは、宿題の量が数学の成績に与える影響に関する回帰係数 β_0 および β_1 を、 n 個のデータすべてを用いて推定する。推定された β_0, β_1 はすべての中学に共通の値である。

一方、 e_i は宿題の量で説明しきれない残差すべてであり、IQや生活態度など個人要因によるばらつきと、教育方法など学校要因によるばらつきをすべて含んでいる。

また、同じ学校に所属する生徒は相互にある程度相関しているはずであるから、このモデルにおける仮定「 n 個のサンプル(データ)はすべて互いに独立である」は厳密には成り立たない。このような級内相関が存在すると、有意性検定が甘くなり、第一種の過誤が生じる確率が増加することが知られている。

一方、各学校の効果を調べるために、それぞれの学校毎に回帰直線、つまり、切片 β_{0j} と傾き β_{1j} を推定する方法も考えられる($j = 1, \dots, m$ はそれぞれの学校を表す)。しかし、そのような分析から得られる結果は各学校に対する推定結果であり、本来の目的であった京都市内の中学に在学する生徒に対する推測に用いることは出来ない。また、各学校ごとのサンプル数が小さい場合には推定誤差が非常に大きくなってしまう。

通常の分散分析や回帰分析では個人が無作為抽出されると仮定するのに対して、本稿で紹介する線形混合モデルは、データがもつ階層構造あるいはグルー

ブ構造を直接モデル化する。そのため、階層線形モデルあるいはマルチレベルモデルと呼ばれることもある。個人差によるばらつきと、個人が所属する上位レベルのグループの違いによるデータのばらつき（たとえば、学校の違いに起因する個人の成績のばらつき）を同時に、区別して評価することができる。同一グループ内のデータ同士は他のグループのデータとよりも類似性している可能性があることを認めた上で、個人要因（IQや親の経済力など）と個人が所属するグループの要因（学校の種類や教育方針など）が個人の結果に与える影響の大きさを評価することができるのである。

線形混合モデルは、医学統計分野および教育分野で先行している分析方法であるが、その他の社会科学分野にも広く適用可能であり、かつ、学問上のニーズは高いと思われる。実際、国内での応用事例は少ないものの、すでに海外ではさまざまな研究分野で利用され、階層構造あるいはグループ構造を持つデータを解析する一般的なツールとして浸透してきている。Kreft and de Leeuw(2006)は、すでに応用が始まっているか、今後の応用が特に期待される分野として、学校の影響力に関する研究、企業における労働者の収入、薬物防止に関する研究、心理療法、成長曲線分析(反復測定データの分析)、地理的情報システム、メタ分析、双子と家族に関する研究を挙げている。

国内での研究例としては、小宮山（2000）、市田ら(2005)、河本（2006）、久保川(2006) などがある。また、科学研究費補助金採択課題・成果概要データベース（KAKEN）で「マルチレベル分析」をキーワードに検索すると、ソーシャル・キャピタルや公衆衛生に関する研究など7件の研究課題が該当する（2007年9月21日現在）。

以下、本稿では、線形混合モデルにおける基礎概念と基本モデルを概説し、教育分野でのいくつかの分析事例を紹介する。2章では最も基礎的な線形混合モデルであるランダム切片モデルを、3章ではランダム係数モデルを紹介する。4章では、反復測定データ（同一の被験者を長期間にわたり追跡する縦断的研究など）への応用について述べ、これが線形混合モデルの有望な適用分野の1つであることを示す。さらに5章では混合モデルに必要なデータ数の目安につ

いて、6章では混合モデルを扱うためのソフトウェアや情報源を紹介する。

2 ランダム切片モデル

まず、最も単純な混合モデルとして、切片だけにランダム効果を導入するランダム切片モデルを考えよう。

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + e_{ij} \quad (3a)$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad (3b)$$

$$u_{0j} \sim N(0, \sigma_{0u}^2) \quad (3c)$$

x_{ij} は説明変数、 y_{ij} は目的変数で、添え字*i*は個人を、添え字*j*は個人が属するグループ(集団)を表している。一見すると通常の回帰式に見えるが、このモデルでは u_{0j} があらたに確率変数として加わり、切片 $\beta_0 + u_{0j}$ はグループ毎に異なる値をとる。回帰直線の傾き β_1 は従来どおりグループ間で共通の定数である。

全データに共通の定数である β_0 と β_1 をモデルの固定効果、個人やグループ間で変動する e_{ij} と u_{0j} をランダム効果(変量効果)と呼ぶ。グループレベルのランダム効果 u_{0j} は、データ全体から計算される切片の値 β_0 からのグループ*j*の切片の値の偏差を表し、個人レベルのランダム効果 e_{ij} はそれぞれのグループレベルにおける平均的な効果 $\beta_0 + \beta_1 x_{ij} + u_{0j}$ から個人の観測値 y_{ij} への偏差を表している。 u_{0j} と e_{ij} は互いに独立である。

固定効果とランダム効果が混合しているモデルなので混合モデルと呼ばれる。分析対象となっているグループは、グループ母集団からのランダムサンプルとしてたまたま選択されたものだと仮定することによって、グループ要因による y_{ij} のばらつきの大きさの指標 σ_{0u} を推定する。

このランダム切片モデルは、説明変数で説明しきれない誤差項を、グループレベルのランダム効果 u_{0j} と個人レベルのランダム効果 e_{ij} という異なるレベル(階層)の分散成分に分けてモデル化しているため、分散成分分析とも呼ばれる。

2.1 級内相関

グループレベルの分散 σ_{0u}^2 および個人レベルの分散 σ_e^2 と、通常の回帰モデルにおけるランダム誤差の分散 σ^2 との間には、次のような関係がある。

$$\sigma^2 = \sigma_{0u}^2 + \sigma_e^2$$

ランダム切片モデルの場合、同じグループ内のデータ同士の相関係数（級内相関）は次のような式で表される。

$$\frac{\sigma_{0u}^2}{\sigma_{0u}^2 + \sigma_e^2}$$

これは、誤差変動全体のなかでグループの違いが原因として生じる誤差変動の割合（VPC, variance partition coefficient）であるから、グループの重要性あるいは効果の程度を表していると解釈することもできる（H. Goldstein *et al.*, 2002）。

級内相関が存在する場合、第一種の過誤の確率が増加し、伝統的な線形モデルの有意性検定が甘くなることは既に述べた。Kreft and de Leeuw(2006)では、大きなグループ $n_j = 100$ においては、0.01という小さな級内相関であっても、仮定された有意水準は $\alpha = 0.05$ から0.17に第一種の過誤率が増大し、小さなグループ（ $n_j = 10$ ）においても、級内相関が大きい（0.2）場合には有意水準 $\alpha = 0.05$ は0.28という観測 α 水準になるという実験結果が紹介されている。

図1は、ロンドンの学校での入学前読解テストと卒業時試験スコアとの関係を、65の学校に所属する4059人の生徒についてプロットしたものである。データはすべて正規化されている。このデータは、

<http://www.cmm.bristol.ac.uk/learning-training/multilevel-m-software/data-rev.shtml>

からダウンロードできる（以下では、ロンドンデータと呼ぶことにする。このデータについての詳細は、Goldstein *et al.* (1993) を参照のこと）。入学前の読解テストと卒業時試験スコアには正の相関関係がみとめられるが、その一方で、かなりのばらつきが存在する様子がみてとれる。なお、図中の直線は、通

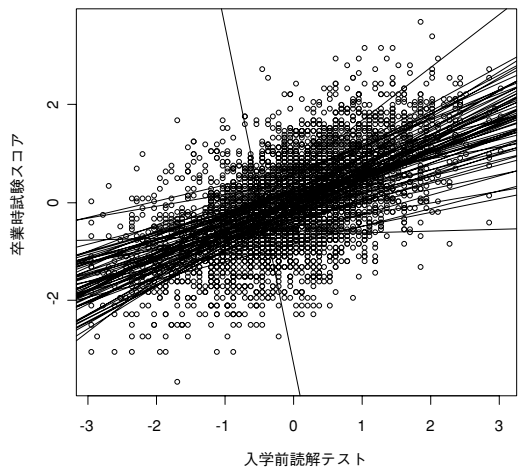


図1 ロンドンの65校における入学前読解テストと卒業時試験スコア

表1 ロンドンデータに対するランダム切片モデルの推定結果

固定効果	(95%信頼区間)	ランダム効果	(95%信頼区間)
$\hat{\beta}_0$	0.002 [-0.077,0.081]	$\hat{\sigma}_{0u}$	0.306 [0.251,0.374]
$\hat{\beta}_1$	0.563 [0.539,0.588]	$\hat{\sigma}_e$	0.752 [0.736,0.769]

常の線形回帰モデルに基づいて、学校毎に推定された回帰直線である。

図2に各学校の回帰直線の切片 β_{0j} (intercept)と傾き β_{1j} (standlrt)の95%信頼区間を示す。切片、傾きともに学校間でのばらつきが大きく、かつ、各学校のサンプルサイズによって、信頼区間の幅が大きく変わることがわかる。また、全く他校と異なる傾向を示す学校が1校あり(図中の48番)、不適切なデータが混在しているらしいことがわかる。

このデータに対してランダム切片モデルを当てはめた結果が表1である。学校レベルのランダム効果の標準偏差は $\hat{\sigma}_{0u} = 0.306$ 、個人レベルのランダム効果、つまり残差の標準偏差は $\hat{\sigma}_e = 0.752$ となり、学校レベルのばらつきもかなりあるという推定結果となった。級内相関を計算すると0.144となり、このデータ

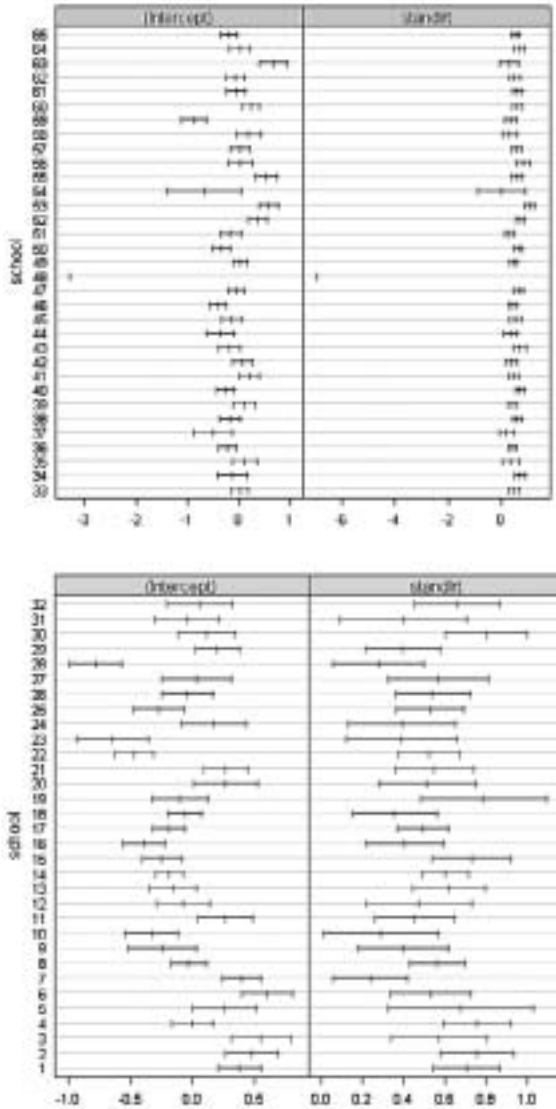


図2 65の学校毎に推定された回帰直線の切片 (intercept) と傾き (standlrt) の95%信頼区間。1番から32番までの図と、33番以降の図では、スケールが異なることに注意。

に対して通常の回帰分析を行うと有意性検定が甘くなることが予想される。

2.2 固定効果とランダム効果

グループの効果を評価する別の方法として、共分散分析 (ANCOVA)がある。回帰モデルにダミー変数を導入し、各グループの効果を固定効果として比較するものである。しかし、 m 個のグループを分析対象とする場合、ANCOVAモデルでは $(m-1)$ 個のパラメータを新たに導入する必要がある(ランダム切片モデルの場合は、グループ数に限らず u_{0j} のみ)。したがって、通常は実験計画の枠組みで、いくつかのグループ間の比較をするのに用いられる(グループの数は多くない)。また、得られた推定結果は分析したグループに対する結果であり、グループが属する母集団に対する推測は行えない。

それに対して、ランダム効果を考える場合には、興味の対象はグループが属する母集団であり、グループ母集団からのランダムサンプルとして各グループのデータを利用するのである。

2.3 ランダム効果の予測

ランダム効果のない通常の線形モデルでは、すべての効果はモデルパラメータとしてモデル式に組み込まれている。たとえば、各グループ毎に線形回帰モデルを考える場合、次のようになる。

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij}, \quad j = 1, 2, \dots, m$$

観測データが与えられれば、パラメータ β_{0j} 、 β_{1j} の値を推定することができる。以下では、この方法でグループ毎に固定効果として計算された推定値のことをOLS (Ordinary Least Square) 解と呼ぶことにする。

一方、ランダム効果を含む混合モデル (3) では、切片のランダム効果 u_{0j} はパラメータではなく、平均が0、分散が σ_{0u}^2 の正規分布に従う確率変数である。その期待値は $E[u_{0j}] = 0$ であり、この値に興味があるわけではない。 $E[u_{0j}] = 0$ というのは、各グループ独自の効果(確率変数 u_{0j} の実現値)を学校間で平均すると0になるというモデルの仮定そのものである。関心があるのは、グルー

グループ毎に割り当てられる確率変数 u_{0j} の実現値の予測（推定）である。混合モデルでは、経験ベイズ最尤法を用い、事前分布 $u_{0j} \sim N(0, \sigma_{0u}^2)$ に対して各グループの観測データが与えられた場合の u_{0j} の事後平均を計算し、それを u_{0j} の実現値に対する予測値(推定値)としている。ランダム切片モデル (3) の場合には、その値は次のような式で表される（久保川 2006）。

$$\hat{u}_{0j} = \frac{\rho n_j}{1 + \rho n_j} (\bar{y}_j - (\beta_0 + \beta_1 \bar{x}_j))$$

ただし、 \bar{x}_j 、 \bar{y}_j は、グループ j 内での x_{ij} 、 y_{ij} の標本平均、 $\rho = \hat{\sigma}_{0u}^2 / \hat{\sigma}_e^2$ である。グループ内のデータ数 n_j が多いほど、 \hat{u}_{0j} のグループ毎の回帰モデルによるOLS解に近づく。また、データ数 n_j が少ない場合は、 \hat{u}_{0j} の値は小さくなり、全体平均に近づいた推定となる。つまり、平均にかなりの程度「縮約(shrinkage)」された推定値となる。データ数が少ない場合、外れ値の影響を受けやすく推定値が歪みやすいという問題があるが、混合モデルの場合は、そのような外れ値の影響を減じることができる。また、級内相関が小さいほど、 ρ の値は小さくなるから、やはりランダム効果 u_{0j} は小さく評価される。

このように、混合モデルにおける各グループに対するランダム効果の予測値は、それぞれのグループに対するOLS解をグループ内のデータ数に応じて全体解に縮約したものである。したがって、将来予測や推論に適した信頼性の高い予測値が得られる反面、データ数が少ないグループに対する推定値はかなりの程度縮約されるため、現状の記述という観点からは現実的ではない（Kreft and de Leeuw, 2006）。

マイノリティに関する研究や家族や双子に関する研究など、グループ毎に十分なサンプル数を確保することが難しいデータを扱う場合、各グループの推定値はかなりの程度縮約される。しかし、多くのグループからのデータをプールし、他のグループのデータから「説得力の借用 (borrowing of strength)」をすることによって、伝統的な統計手法では行えなかった統計的推測が行えるのである。

3 ランダム係数モデル

次に、回帰係数 β_1 にもランダム効果を導入したランダム係数モデルを考えよう。

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij}, \quad \beta_{0j} = \beta_0 + u_{0j}, \quad \beta_{1j} = \beta_1 + u_{1j} \quad (4a)$$

$$u_{0j} \sim N(0, \sigma_{0u}^2), \quad u_{1j} \sim N(0, \sigma_{1u}^2), \quad \text{cov}(u_{0j}, u_{1j}) = \sigma_{u01} \quad (4b)$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad (4c)$$

このモデルでは、係数 u_{0j} 、 u_{1j} がグループレベルのランダム変数である。切片に加えて回帰直線の傾きもグループ毎に異なる値をとる。また、2つのランダム効果 u_{0j} と u_{1j} の共分散を σ_{u01} とする。つまり、混合モデルではランダム効果に関する相関構造を指定できる。このモデルの場合は、共分散 σ_{u01} が正の場合、切片が小さければ係数も小さく、切片が大きければ係数も大きくなる傾向があり、 σ_{u01} が負の場合は、切片が小さければ係数が大きく、切片が大きければ係数は小さくなる傾向があることになる。

このモデルにおけるグループ間の分散、つまりグループの違いによる目的変数 y_{ij} のばらつきの大きさは、

$$\text{Var}[u_{0j} + u_{1j}x_{ij}] = \sigma_{0u}^2 + 2\sigma_{u01}x_{ij} + \sigma_{u1}^2x_{ij}^2 \quad (5)$$

となり、説明変数 x_{ij} の2次関数になる。さらに、2.1節で紹介したVPCは次のような式で与えられる。

$$\frac{\sigma_{0u}^2 + 2\sigma_{u01}x_{ij} + \sigma_{u1}^2x_{ij}^2}{\sigma_{0u}^2 + 2\sigma_{u01}x_{ij} + \sigma_{u1}^2x_{ij}^2 + \sigma_e^2}$$

また、同じグループ内のデータ x_{i_1j} と x_{i_2j} の級内相関は次のように表される。

$$\frac{\sigma_{0u}^2 + \sigma_{u01}(x_{i_1j} + x_{i_2j}) + \sigma_{u1}^2x_{i_1j}x_{i_2j}}{\sqrt{\sigma_{0u}^2 + 2\sigma_{u01}x_{i_1j} + \sigma_{u1}^2x_{i_1j}^2}\sqrt{\sigma_{0u}^2 + 2\sigma_{u01}x_{i_2j} + \sigma_{u1}^2x_{i_2j}^2}}$$

表2にロンドンデータにランダム係数モデルを当てはめた結果を示す。傾きの固定効果 β_0 の推定値が0.557であるのに対して、学校の違いによる傾きのランダム効果 u_{1j} の標準偏差は0.122である。切片についての固定効果とランダム効

表2 ロンドンデータに対するランダム係数モデルの推定結果

固定効果	(95%信頼区間)	ランダム効果	(95%信頼区間)
$\hat{\beta}_0$	-0.012 [-0.090, 0.067]	$\hat{\sigma}_{0u}$	0.303 [0.249, 0.371]
$\hat{\beta}_1$	0.557 [0.517, 0.596]	$\hat{\sigma}_{1u}$	0.122 [0.090, 0.166]
		$\hat{\sigma}_e$	0.744 [0.728, 0.761]
		$\hat{\sigma}_{u01}$	0.494 [0.149, 0.731]

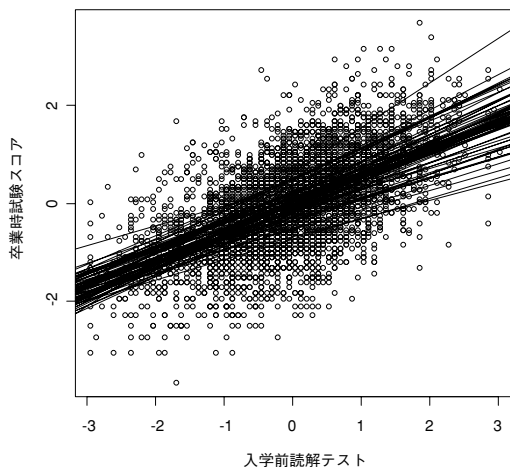


図3 ランダム係数モデルによって推定された65校の回帰直線

果の推定値は、ランダム切片モデルの推定値とほぼ同じである。また、学校レベルの2つのランダム効果（切片と傾き）の相関係数が0.494であるから、ランダム切片が大きい学校ほど、ランダム傾きも大きくなる傾向があることがわかる。図3に、ランダム係数モデルによって推定された各学校の回帰直線を示す。図1のOLS解と比べると、サンプル数の少ない学校に対して「説得力の借用」をすることによって、安定した推定結果が得られることがわかる。また、読解力の優れた生徒ほど卒業時のスコアに対する学校の影響が大きいことがわかる。

図4はロンドンデータの入学前読解テストの得点の関数としてVPCをプロッ

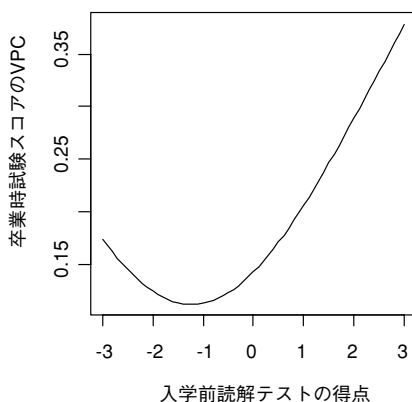


図4 入学前読解テストの得点と卒業時試験スコアのVPC

トしたものである。この図からも、卒業時の成績に対する学校の効果は入学時の読解力の程度によって大きく異なり、とくに読解力が平均以上の生徒に対しては、読解力の優れた生徒であるほど入学する学校の影響が大きいことがわかる (H. Goldstein *et al*, 2002)。また同時に、読解が苦手な生徒にとっても学校の効果は無視できないものであり、読解力の優れた生徒に有効な学校が必ずしも読解の苦手な学生に対しても有効であるとは限らない (図4 参照)。

3.1 モデルの比較

では、同じデータセットに対して複数のモデルを当てはめた場合、モデルの当てはまりの比較はどのように行うのであろうか。混合モデルの場合には赤池情報基準 (AIC)、ベイズ情報基準 (BIC)、対数尤度 (log Likelihood)、逸脱度 (deviance) などの統計量を用い、これらの統計量の値が小さいほどモデルがより適合していると判断する。Kreft and de Leeuw (2006) は、モデルパラメータ数と適合度の関係について、「大ざっぱなルールとして、あるモデルが他のモデルよりも有意に改善されたと結論するには、2つのモデル間の逸脱度の差が最低、推定パラメータ数の差 (自由度の減少分) の2倍は大きいべきである。」と述べている。

表3 2つのモデルの適合度の比較

モデル	df	AIC	BIC	logLikelihood	test	L.Ratio	p値
ランダム切片(1)	4	9376.77	9402.00	-4684.383			
ランダム係数(2)	6	9339.60	9377.45	-4663.80	1 vs 2	41.16	<.0001

表3は、ランダム切片モデルとランダム係数モデルのロンドンデータに対する適合度を尤度比検定により比較した結果である。ランダム係数モデルにすることによってパラメータは2つ増えるが、AICが37程度下がるなどすべての指標で大幅に下がっている。したがって、このデータに対してはランダム係数モデルが統計的により適合しているといえる。

3.2 傾きを結果変数とするアプローチ

ところで、通常の回帰モデルにおいてグループ間の違いを分析する場合、グループ毎に回帰直線のOLS解を求め、それらの推定値をさらにグループレベルの変数を説明変数として回帰分析にかけるという方法が用いられることがある(「2ステップ分析」あるいは「傾きを結果変数とする」アプローチと呼ばれる)。しかし、この方法では、分析の最小単位は個人ではなくグループとなる。グループで集計された回帰係数の値は、サンプル数が異なればその有効性(標準誤差の違いなど)も異なるが、それらは無視され同等に扱われてしまう。また、グループでの集計値を用いて得られた結果から個人の傾向について言及することはできない。このような誤用は生態学的誤謬(ecological fallacy)として知られている。

しかし、その一方で、このアプローチは非常に魅力的である。ランダム係数モデルでは、分析を2ステップに分けることなく、単一のモデルで「傾きを結果変数とする」アプローチを行うことができる。つまり、グループの特性や何らかの社会的文脈を用いて、グループ間の切片や傾きのばらつきを「説明する」ことができるかどうかを評価することができる。

各グループの特性を表すグループレベルの変数 z_j を考えよう。これは、たと

えば、学校に対しては公立かどうか、男女の構成比、入学試験での平均得点（学校のランク）などであり、地域に対しては平均所得や人種構成などである。

ランダム係数モデル（4）において、回帰係数 β_{0j} および β_{1j} がグループレベルの変数 z_j を用いて、次のような式で部分的に説明できるとする。

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u'_{0j}, \quad \beta_{1j} = \gamma_{10} + \gamma_{11} z_j + u'_{1j}$$

これらを（4）に代入して整理すると次の式が得られる。

$$y_{ij} = \gamma_{00} + \gamma_{01} z_j + \gamma_{10} x_{ij} + \gamma_{11} z_j x_{ij} + (u'_{0j} + u'_{1j} x_{ij} + e_{ij})$$

固定効果（ $\gamma_{00} + \gamma_{01} z_j + \gamma_{10} x_{ij} + \gamma_{11} z_j x_{ij}$ ）には、交互作用項 $\gamma_{11} z_j x_{ij}$ が新たに加わっているが、ランダム効果は同じ形のままである（もちろん、それらが意味するものは（4）とは異なる）。

では、実際の分析例として、家族の社会経済的地位が数学の成績にどのように影響を及ぼすかを手短かに示そう（Fox, 2002）。この分析で用いられたデータ（以下、MathAchと呼ぶ）は、統計解析ソフトRのnlmeパッケージに同梱されているため追試することができる。

生徒の家族の社会経済的地位(cses)を x_{ij} 、生徒の数学の成績を y_{ij} とするランダム係数モデル（4）において、ランダム効果（ β_{0j} , β_{1j} ）を各学校の生徒の家族の社会経済的地位の平均 z_{1j} を用いて説明するモデル（モデル1）：

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_{1j} + u'_{0j}, \quad \beta_{1j} = \gamma_{10} + \gamma_{11} z_{1j} + u'_{1j} \quad (8)$$

と、さらに各学校のタイプ（公立か私立か） z_{2j} もランダム効果の説明変数とするモデル（モデル2）：

$$\beta_{0j} = \gamma_{00} + \gamma_{01} z_{1j} + \gamma_{02} z_{2j} + u'_{0j}, \quad \beta_{1j} = \gamma_{10} + \gamma_{11} z_{1j} + \gamma_{12} z_{2j} + u'_{1j} \quad (9)$$

を考える。ただし、公立学校に対して $z_{2j} = 0$ 、私立学校に対して $z_{2j} = 1$ とする。この2つのモデルのデータMathAchへの当てはまりを比較した結果が表4である。明らかにモデル2の方がよい。

次に、モデル2におけるモデルパラメータの推定結果を示す（表5）。まず、家族の社会経済的地位だけを説明変数とするランダム係数モデル（モデル0）と比べてみよう。モデル0では、 $\hat{\sigma}_{0u} = 2.95$ 、 $\hat{\sigma}_{1u} = 0.833$ 、 $\hat{\sigma}_e = 6.06$ という結果になる。モデル2では $\hat{\sigma}_{0u} = 1.54$ 、 $\hat{\sigma}_{1u} = 0.32$ 、 $\hat{\sigma}_e = 6.06$ であり、切片や傾

表4 データMathAchに対するモデル1とモデル2の適合度の比較

モデル	df	AIC	BIC	logLikelihood	test	L.Ratio	p値
モデル1	8	46573.30	46628.34	-23278.65			
モデル2	10	46523.66	46592.45	-23251.83	1 vs 2	53.64	<.0001

表5 MathAchに対するモデル2の推定結果

固定効果	(95%信頼区間)	ランダム効果	(95%信頼区間)
$\hat{\gamma}_{00}$	12.13 [11.74, 12.52]	$\hat{\sigma}_{0u}$	1.54 [1.32, 1.80]
$\hat{\gamma}_{01}$	5.33 [4.60, 6.06]	$\hat{\sigma}_{1u}$	0.32 [0.06, 1.59]
$\hat{\gamma}_{10}$	2.95 [2.64, 3.25]	$\hat{\sigma}_e$	6.06 [5.96, 6.16]
$\hat{\gamma}_{02}$	1.23 [0.62, 1.83]	$\hat{\sigma}_{u01}$	0.39 [-0.51, 0.88]
$\hat{\gamma}_{11}$	1.04 [0.45, 1.63]		
$\hat{\gamma}_{12}$	-1.64 [-2.11, -1.17]		

きの学校間のばらつきのかかなりの部分を学校レベルの変数で説明できることがわかる。

さて、モデル2のみで説明変数として採用した学校のタイプを表す変数 z_{2j} の効果はどのようなものかを見てみよう。 $\hat{\gamma}_{00}=12.13$ は公立学校における数学のテストの平均値に相当し、私立学校における平均値はさらに $\hat{\gamma}_{02}=1.23$ を加えたものになる。また、公立学校における家族の社会経済的地位の係数は $\hat{\gamma}_{10}=2.95$ であるが、私立学校においては、その係数は $\hat{\gamma}_{10}+\hat{\gamma}_{12}=1.31$ となる。したがって、私立学校のほうが公立学校と比べて、数学の平均点が高く、かつ、生徒の家庭の社会経済的地位が数学の成績に与える影響が小さいことがわかる。公立と比べると、私立の方が平等主義的な教育を実践していると解釈できるかもしれない。

4 経時データの解析

ある個体について、時間を追って複数回データを測定することを反復測定と

いう。ランダム係数モデルは、このような反復測定データ(経時データ、縦断データと呼ぶこともある)の分析に応用できる。最も典型的な反復測定データの例として、個人に対して繰り返し測定を行う場合を考える。この場合、各個人のデータ集合が1つのグループを構成し、添え字 j が個人を表す。説明変数 x_{ij} は各個人の i 番目の測定時間に対応する変数となる。実験研究の場合には、実験開始からの経過時間であるだろうし、発達に関する研究の場合には、測定時の被験者の年齢と考えてもよい。

(4) で定義したランダム係数モデルにおいて、時間経過を表す変数 t を説明変数 x_{ij} とする場合を考える。すると、(5) より t 時間経過後の被験者間の分散は次のようになる。

$$\sigma_{0u}^2 + 2\sigma_{u01}t + \sigma_{u1j}^2 t^2 + \sigma_e^2$$

したがって、被験者間の変動の大きさは時間 t の関数になり、時間と共に増大する(共分散 σ_{01} が負の場合は一定時間経過後に増大する)。

一方、同一被験者内の異なる2時点 t_1 、 t_2 での測定データはお互いに相関していると考えられる。特に、非常に近い時間間隔で測定されたデータは似たものになりがちであろう。したがって、反復測定データを分析する場合、個人レベルのランダム誤差 e_{ij} を独立同一な分布からのランダムサンプルだとする仮定は必ずしも適切ではない。混合モデルでは、被験者内ランダム誤差 e_{ij} に対して自己相関構造を指定するなど、さまざまな共分散構造を指定することができる。

伝統的な反復測定ANOVAやSEMの1つとしての潜在曲線モデルと違い、個人毎のデータ数や観測のタイミングが同じでなくてもよい。つまり、アンバランスデータに対応できるという非常に大きなメリットがある。つまり、反復測定データに対して、ランダム係数モデルは非常に自然で便利なモデルを提供する。また、目的変数は必ずしも時間(年齢) t について線形にする必要はなく、たとえば、 t と t^2 を説明変数とすることもできる(Steele and Goldstein, 2007)。

反復測定データへの混合モデルの適用や既存の分析法との比較については、

数多くの文献が存在している。詳しくは、例えば、千野（2003）、Van der Leeden(1998)およびそれらの引用文献を参照のこと。

5 有効な推定に必要なデータ数の目安

混合モデルで有効な推定結果を得るためには、全体としてのサンプル数だけでなく、グループ数と各グループ内のサンプル数の両方が一定数以上必要である。特に、グループ数が少ない場合にはランダム効果が過小評価されたり、大きな標準誤差を持ったりする。また、3.2節で紹介したようなクロス水準の交互作用を加えると多重共線性が生じやすく、モデルが不安定化しやすい。

実際、ロンドンデータに対するランダム係数モデルの推定において、 $\hat{\sigma}_{1u}$ の95%信頼区間は[0.090,0.166]であるのに対して、MathAchデータにおける $\hat{\sigma}_{1u}$ の95%信頼区間は[0.064,1.59]であり、MathAchデータでは標準誤差がかなり大きくなる。

クロス水準の交互作用を組み込んだ混合モデルの場合、最低20以上のグループが必要であるといわれているが、シミュレーション研究からは、次のような目安が提示されている（Kreft and de Leeuw, 2006）。

- ・グループ数150の場合、グループ当たりのデータ数は5
- ・グループ数60の場合、グループ当たりのデータ数は25
- ・グループ数30の場合、グループ当たりのデータ数は30

なお、十分なデータ数が確保できる場合には、通常の回帰モデルと混合モデルは固定効果に対してはほぼ同じ推定値を出力する。

6 ソフトウェアおよび情報源

混合モデルでは、最尤推定法を用いて固定効果やランダム効果の共分散構造などすべての推定パラメータを同時に推定する。かなり条件の悪い非線形最適化の計算が必要であるが、コンピュータアルゴリズムや専用ソフトウェアが開

発され、一般の統計ユーザーにとってもすでに実用的なものとなっている。専用パッケージにはMLwiN、HLM、SuperMix、aMLなどがある（SuperMixは2007年9月にリリース予定）。aMLは無料、HLMとSuperMixはstudent editionが無料、MLwiNはtraining versionが無料であり、下記のWebサイトでそれぞれ詳しいマニュアル、オンラインテキスト、適用事例、サンプルデータ、関連文献など豊富な情報が提供されている。

- ・ **MLWin** Centre for Multilevel Modelling(CMM)

<http://www.cmm.bristol.ac.uk/>

Teaching Resources And Materials for Social Scientists(TRAMSS)

<http://tramss.data-archive.ac.uk>

- ・ **HLM** Scientific Software International, Inc (SSI)

<http://www.ssicentral.com/hlm/index.html>

- ・ **SuperMix** Scientific Software International, Inc (SSI)

<http://www.ssicentral.com/supermix/index.html>

- ・ **aML** aML Multiprocess Multilevel Modeling

<http://www.applied-ml.com/>

さらに、フリーの統計解析ソフトRには混合モデルを扱うためのパッケージとしてnlmeやlme4が公開されており、Rユーザーなら簡単に混合モデルを利用できる環境が用意されている。また、具体的な分析手法についての解説もある(Fox, 2002; 林, 2004; Faraway, 2006)。本研究ではRを用いている。

SPSSのAdvanced Modelオプションにも線形混合モデルが含まれている。他のパッケージが一般化線形混合モデル（GLMM）も分析できるなど、さまざまなオプションが用意されているのに対して、SPSSでは正規性を仮定するモデルのみが対象であること、誤差構造の指定が限定されること、出力される統計量に制限があることなど、いくつかの欠点が指摘されている(A. H. Leyland 2004)。しかし、SPSSは日本語環境で利用可能であり、SPSSに慣れたユーザーは手軽に利用できる。

MLwiN、HLM、R、SPSSについては、日本語での簡単な解説および操作説

明が『基礎から学べるマルチレベルモデル』（小野寺編訳，2006）に掲載され、国内ユーザーにとっても身近なものになりつつある。また、上で紹介した Centre for Multilevel Modelling の Web サイトに、混合モデル（マルチレベルモデル）を扱うさまざまなソフトウェアの詳細なレビューと推定結果や計算時間の比較結果が公開されており、混合モデルで分析を試みる場合に大変参考になる。

最後に紹介する winBugs は、応用統計学者向けに開発された、マルコフ連鎖モンテカルロ法（MCMC）を用いてベイズ推定を行うフリーの専用プログラムである。他のパッケージと異なり、パラメータ推定法が最尤法ではなく確率的な MCMC である。確率的なシミュレーションを行うため、計算時間を要するという欠点はあるが、より正確な推定値を得ることができる。また、解析やモデル構築を Windows インターフェースやグラフィカルインターフェースを介して行うことができ、柔軟なモデル構築に対応できるというメリットもある。さらに、winBugs は単一のプログラムであるため、他のソフトウェアから呼び出して使うことができる。winBugs の Web サイトには、R、STAT、SAS、Matlab、Excel などから呼び出すためのツールが公開されている。

・ **winBugs** The BUGS project – WinBUGS

<http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml>

おわりに

無作為抽出ではない集団化されたデータやスパースなデータを扱う場合、あるいは、経時データを扱う場合、混合モデルは大きな可能性を秘めた重要な分析手法の一つである。実際、混合モデルの方法論的な文献および応用に関する文献は急速に増えてきている。

本稿では、ランダム効果を組み込むモデルとして最も基本的な線形混合モデルを紹介したが、この他にも、ロジスティック回帰やポアソン回帰などの一般化線形モデルにランダム効果を取り入れた一般化線形混合モデル（GLMM）、

因子分析や共分散構造分析にランダム効果を仮定する混合モデルなど、さまざまなタイプの混合モデルが提案され、すでに多くの解析パッケージに組み込まれている。

混合モデルはグループ(集団)や社会的文脈が個人に与える影響の大きさを分散で表すモデルであり、個々の集団やグループの特徴を詳細に記述するものではない。したがって、学校のような現実のグループに関する意思決定(たとえば学校の順位付け)に利用するのはあまり適当でないが、集団や社会的文脈などさまざまな分野における理論を実証するための強力なツールとなりうる。今後、多くの研究分野でこの手法が活用されることを期待したい。

参考文献

- [1] Faraway, J. J. (2006) Extending the Linear Model with R – Generalized Linear, Mixed Effects and Nonparametric Regression Models. Chapman & Hall/CRC.
- [2] Fox, J. (2002) Linear Mixed Models, Appendix to An R and S-PLUS Companion to Applied Regression.
CRAN(The Comprehensive R Archive Network)のホームページから入手可能。
<http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-mixed-models.pdf>.
- [3] Goldstein, H., Rasbash J., Yang M., Woodhouse G., Pan H., Nuttall D. and Thomas S., (1993) A multilevel analysis of school examination results. Oxford Review of Education, Vol.19, 425-433.
- [4] Goldstein, H. (1995) Multilevel Statistical Models. New York: Halsted.
- [5] Goldstein, H., Browne, W. and Rasbash, J. (2002) Partitioning Variation in Multilevel Models. Understanding Statistics, Vol.1, No.4, 223-231.
- [6] Goldstein, H. (2007) Becoming familiar with multilevel modelling. Significance, Vol.4, No.3, 133-135.

文献[4][5][6]は、下記のH.Goldsteinのホームページから入手可能。

http://www.cmm.bristol.ac.uk/team/HG_Personal/index.shtml

- [7] 林啓一 (2004) R でマルチレベルモデリング. 岡田昌史 (編) 『The R Book —データ解析環境Rの活用事例集』 九天社, 340-363.
- [8] 市田行信, 吉川郷主, 平井寛, 近藤克則, 小林槇太郎 (2005) マルチレベル分析による高齢者とソーシャルキャピタルに関する研究 — 知多半島28校区に居住する高齢者9,248人のデータから—. 農村計画学会誌, Vol.24, 277-282.
- [9] 河本幸子(2006) 岡山市内における3歳児う蝕有病者率の地域格差について — マルチレベル分析による検討 —. 口腔衛生会誌, Vol.56, 660-664.
- [10] 小宮山智志 (2004) 階層線形モデルによる"地域不公平感"の分析. 新潟国際情報大学情報文化学部紀要, Vol.7, 161-178.
- [11] Kreft, I. and Yoon, B. (1994) Are multilevel techniques necessary? An attempt at demystification. Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8, 1994),
オンラインデータベース ERIC(<http://eric.ed.gov/>)から入手可能。
- [12] Kreft, I. and de Leeuw, J. (2006) 『基礎から学ぶマルチレベルモデル—入り組んだ文脈から新たな理論を創出するための統計手法—』. 小野寺孝義(編訳), ナカニシヤ出版. (Kreft, I. and de Leeuw, J. (1998) 『Introducing Multilevel Modeling』, Sage Publications Ltd)
- [13] 久保川達也(2006) 線形混合モデルと小地域の推定. 応用統計学, Vol.35, No.3, 137-160.
- [14] Van der Leeden, R.V. (1998) Multilevel analysis of repeated measures data. Quality & Quantity, Vol.32, 15-29.
- [15] McCulloch, C. E. and Searle, S.R. (2001) Generalized, Linear, and Mixed Models. John Wiley & Sons, Inc.
- [16] 村澤昌崇 (2005) 高等教育における計量分析手法の応用(その1) — マルチレベル分析 —. 広島大学高等教育研究開発センター大学論集, Vol.37, 309-327.

- [17] 千野直仁 (2003) 反復測定データの分析とその周辺 (測定・評価部門) . 教育心理学年報, Vol.42, 107-118.
- [18] Snijders, T. A. B. (2003) Multilevel Analysis. in M. Lewis-Beck, A.E. Bryman, and T.F. Liao (eds.), The SAGE Encyclopedia of Social Science Research Methods, Vol.II, Sage, 673-677.
- [19] Steele, F. and Goldstein, H. (2007) Multilevel models in psychometrics. Handbook of Statistics, Vol.26, 401-419.